DOCUMENT RESUME

ED 298 504 CS 211 470

AUTHOR Formsma, Rene
TITLE Literature on Evaluation of Handwriting.
PUB DATE 20 Jul 88
NOTE 30p.; Exit project, Indiana University at South Bend.
PUB TYPE Reference Materials - Bibliographies (131) -- Information Analyses (070) -- Dissertations/Theses - Practicum Papers (043)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Annotated Bibliographies; Elementary Education; Evaluation Methods; *Handwriting; *Writing Instruction; *Writing Research; *Writing Skills
IDENTIFIERS Legibility

ABSTRACT

This project is a historical and topical search into handwriting evaluation. It includes a statement of the problem; a justification; definitions; a bibliography listed in historical order with annotations on 36 books and articles; and a summary and evaluation of the literature. The conclusion of the paper discusses the importance of being aware of evaluation techniques and tools that can be used to judge students' handwriting skills. The paper closes with three recommendations for teachers: (1) choose a handwriting evaluation scale, method, or tool and use it consistently; (2) have students self check their handwriting by using transparent overlays to evaluate daily progress; and (3) use an evaluation plan as a tool which lends itself to remediation. Thirty-five references are appended. (MS)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED298504

Literature on Evaluation of Handwriting

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Rene A Formsma

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

by

Rene Formsma

Education E591

Charles DuVall

July 20, 1988

CS211470

TABLE OF CONTENTS

Introduction 1

Statement of the Problem 2

Justification 2

Outline 2

Definitions 3

Annotations 4

Summary 14

Conclusion 21

Recommendations 23

Bibliography 24

Student Background 27

Introduction

Evaluation of handwriting by using statistic based methods are not a new idea. The need for a systematic approach started with Thorndike in the early 1900's. Throughout the decade various research has been done to perfect an evaluation tool. It is the authors intent to make a historic and topical search into handwriting evaluation.

Statement of the Problem

In order to help students succeed in handwriting skills, educators need to be aware of objective ways to evaluate handwriting.

Justification

Teachers need to be able to evaluate handwriting samples with reasonable consistency, validity and reliability. Graham,(1986) states, "The evaluation of handwriting products are often imprecise and based on subjective criteria." There are many criteria to evaluate handwriting that educators must be familiar with in order to be effective in the teaching of handwriting. One criteria may be legibility which is used as the basis for the "Ayres Scale". Factors such as uniformity of height, slant, quality of line, appearance and spacing are characteristics in judging excellence in writing as seen by Freeman.(1915)

It is through an understanding of planned handwriting evaluation that students skills will be enhanced.

Outline

The proposed study will be organized into these parts: an overview which includes definitions, literature on techniques of evaluating handwriting placed in historical order with some shifting for ease of reporting topics. In closing, there will be a summary, conclusion and recommendations.

Definitions

Characteristic for judging excellence in handwriting-(Freeman)-uniformity of slant, uniformity of alignment, quality of line, letter formation and legibility.

Correct/Incorrect Method-evaluation system that uses transparent overlays to give evaluation scores to letters by degrees of formation.

forced rank-evaluation that uses four criteria, overall appearance, smoothness of strokes, closure, and stroke level with respect to lines on the paper.

general merit-average judgement of competent people-quality, goodness.

global rating-letters in a work are rated on a scale from 0-4. 0 is an illegible letter and 4 indicates a well formed letter.

handwriting scale-samples of handwriting usually ranked in order highest to lowest based on some set criteria. The function of the scale is to be a standard by which students handwriting can be judged.

holistic rating system-Likert type scale which gives scores of 0-4 based on a set criteria.

legibility-(Ayres)-time needed to easily read a handwriting sample.

legibility-(Freeman)-based on the space between letters words and lines in handwriting.

letter formation-way in which letters are made.

quality of line-way to evaluate handwriting. Based on lines regularity, smoothness, evenness or lack of it.

self instruction-use of self verbalization by students to train themselves in handwriting.

transparent overlay-plastic overlays that have ellipse-like boundaries printed on them to evaluate letters.

Annotations Listed in Historical Order

Thorndike, Edward L., Handwriting. Teachers College, Columbia University, New York, 1912.

Thorndike developed a handwriting scale that gave a basis for judging handwriting samples on, "general merit." He contends his scale may stimulate others to go ahead and do more research in this area. His scale consists of quality samples of handwriting ranked from a low of four to a high of 18. These samples were ranked on the "average judgement" of competent people. Criticisms of his scale include that only one style of writing is represented and not all samples are of actual childrens work and perhaps legibility is a better criteria than general merit in evaluating handwriting. Thorndike included a scale for, "Adult Womens Writing" also speed and quality relationships for seven school systems.

Ayres, Leonard P., A Scale for Measuring the Quality of Handwriting of School Children. No. 113, Russell Sage Foundation. 1912.

This study details a handwriting scale that was produced to measure legibility of upper elementary school children. The study used 1,578 samples of handwriting which was collected from 40 schools in 38 states. Each sample was judged by ten people according to the time needed to read it. The scale was made up of actual students samples. This scale is on a vertical 9"x36" paper and gives three samples of equally legible writings with varied slants. This tool ranks the samples 10, 20, 30... 90 which can easily be converted to grades by a classroom teacher. Factors which were found to effect legibility are spacing between words, spacing between lines and slant.

Pitner, Rudolf, A Comparison of the Ayres and Thorndike Handwriting Scales. Journal of Educational Psychology, 1914, 5, 525-536.

Two studies were done to compare the Ayres scale to the Thorndike scale. The problem was to see which scale would yield the most uniform results. The first study used 24 samples and the second used 273. In both cases, the ranking of samples by each scale was significantly similar. The scales were compared by deviation and by this comparison, Thorndike's scale gave more reliable and uniform results.

Kelley, Truman Lee, Comparable Measures. Journal of Educational Psychology, 1915, 6, 589-595.

This author disputes the finding of Pitner.(1914) This article refers to the comparison of the Thorndike and Ayres scales. The author used his formulas, which reduced the scales to more comparable units. He found the Ayres scale showed slight superiority in uniformity.

King, I., & Johnson, H., The Writing Abilities of the Elementary and Grammar School Pupil of a City School System Measured by the Ayres Scale. Journal of Educational Psychology, 1912, 3, 514-520.

The purpose of this study was to apply the Ayres scale to school age children with reference to age, grade and school variation. The study used 966 samples of handwriting collected from children from first to eighth grade and from six to seventeen years old. The results showed that a single judge can use this scale with a fair amount of consistency. When comparing the results of eight judges, it was impossible to eliminate subjectivity factors. The scale did keep subjective grading to a minimum. The progress of students is rapid between grades one and two or between the age of six and seven. All schools obtained fairly uniform results.

Manuel, Herschel T., The Use of and Objective Scale for Grading Handwriting. The Elementary School Journal, 1915, 15, 269-278.

The purpose of this paper is to address the "variability of grades" assigned to the same writing when using the Ayres scale. The study used experts, as well as people with no training, in the scales use to grade sampled work. The scale method showed no superiority over the school method in ranking papers. The key to reducing deviation among scores is experience as individual deviations were less over a period of time. The significance of the paper is in data gathering and furnishing more information on the problem of "scales" in educational measurement.

Gary, Truman C., The Training of Judgement in the Use of the Ayres Scale for Handwriting. Journal of Educational Psychology, 1915, 6, 85-98.

The effects of the mental process on the use of the Ayres scale were evaluated. Findings show that reliability of the scale may be improved with training. Three judges were trained in the use of the Ayres scale. Samples were evaluated in an experiment for a period of twenty weeks. During the course of the study, weekly discussions on judgements were held. By the sharing of information and practice, variability was reduced. The mental processes involved three steps. The steps were: getting a general impression of the sample, forming general impressions of the samples on the scale and comparing the sample to the scale. It was found that after a period of time, the scale was only used for reference. Training proved to reduce variability to a minimum. The author recommended its use in college preparation courses. Although time was extensive in this study, techniques could be developed to reduce training time.

Freeman, Frank, An Analytical Scale for Judging Handwriting. The Elementary School Journal, 1915, 15, 432-441.

This article details the steps which are necessary to increase reliability of all handwriting scales. He feels it's important to define the criteria for judging and give objective ways to do it. The author makes reference to his scales use of slant, height, letter formation, quality of line and appearance. He gives definitions of these terms and specific means to measure each character. He addressed this problem with "scales of the past." Scales of the past would yield more reliable results if training in their use occurred.

Freeman, Frank Nugent, The Zaner-Bloser Progress Record and Scale in Handwriting, Zaner-Bloser Co., Columbus, Ohio, 1929.

This is a progress record for, grade eight, handwriting. Freeman scale is used. The purpose of the scale is given, as well as specific ways to judge the five areas of legibility. There is a guide for measuring speed. Samples of students work and grade approximations are included.

Starch, Daniel, Educational Research and Statistics: A Scale for Measuring Handwriting. School and Society, 1919, 9, 155-158.

An assortment of issues are presented dealing with the Thorndike and Ayres Scales. The purpose of this article is to examine both scales to see if a new means of evaluation was necessary. A review of the Thorndike scale found it to be an impractical tool, as its size is awkward and samples of different slants of handwriting are lacking for many quality rankings. One-hundred judges rechecked the values assigned to Thorndike's samples, which were originally judged by 30-40 people. Significant differences in rankings were found. Little criticism was given to Ayres scale, as far as practicality goes. Judges question Ayres legibility reading rate when the scale is compared for general appearance. One-hundred judges used Ayres scale and general merit to recheck the scales rankings. Marked inequalities in the distance of value between steps occurred.

Andersen, Dan W., Handwriting Research: Movement and Quality. Elementary English, 1965, 42, 45-53.

This article reviews the history of the handwriting movement and measurements of its quality from 1910 to present.(1965)

Freeman, Frank N., A New Handwriting Scale. Elementary School Journal, 1959, 59, 218-221.

A handwriting scale was created by Freeman which is scaled by grade. Samples of childrens actual writing were obtained nationwide. The scale includes five specimens for each grade. Samples which are ranked from lowest to highest were obtained from a total of 135,491 samples. The narrowing process occurred in steps. The steps were: 1. sorting by grade, 2. sort into five groups by quality, 3. sort into 3 groups by quality, 4. finally, judge by principle of legibility. The principles given are formation, spacing, alignment, fluency, letter form, uniformity of size and slant. This scale represented the grades of quality throughout the nation.

Feldt, Leonard S., The Reliability of Measures of Handwriting Quality. Journal of Educational Psychology, 1962,53, No. 6, 288-292.

Feldt conducted a study to find reliability data on the Evaluation Scale, Grades 1 and 2, by Freeman.(1959) Five categories of measurement error were reported on using the handwriting samples of approximately 30, grade one and two students. The relationship among the scores of the judges, given by Freeman's scale, were low when compared to achievement test scores. The author felt reliability could be improved by improving training to educators and handwriting scores from testing sessions should be averaged.

Bezzi, Rocky, Standardized Manuscript Scale for Grades 1, 2, and 3. The Journal of Educational Research, 1962, 55, No. 7, 339-340.

Bezzi adds to requirements set up by Freeman.(1959) The requirements cited are: scales for each grade level, measure rates, as well as quality, general criteria should be legibility, samples should be from children in the U.S., use statistical technique which renders validity and reliability, lend itself to reporting progress overlap quality among grades and be cross validated with cursive writing. Bezzi did sampling of 7,212 samples from 130 schools. Using these requirements, Bezzi set up a scale which provides a standardized means by which teachers can evaluate handwriting.

Erlebacher, Adrienne, and Herrick, Virgil E., Quality of Handwriting Today and Yesterday. Elementary School Journal, 1961, 62, 89-93.

A study by Erlebacher and Herrick was done to make a valid assessment of the quality of writing today as it compares to writing in 1912. The samples from the Ayres Scale of Measurement were used and compared with samples of 677 Wisconsin students in 1959. The scores of the Wisconsin samples were judged using Attneavs method of graded dicotomies. Ayres Scale samples were then reassigned values as a means for comparing. Based on these findings, there is a strong indication that samples of 1912 and 1959, do not differ in median legibility. Handwriting does not show signs of deterioration based on this study.

Herrick, V., and Okada, N., The Present Scene: Practices in the Teaching of Handwriting in the U.S.-1960. In V. Herrick (Ed.), New Horizons for Research in Handwriting, Madison: University of Wisconsin Press, 1963.

The survey by Herrick and Okada was done for the purpose of finding out what the instructional practices of handwriting in the United States actually are. The national survey covered these four areas:

1. General nature of instructional program in handwriting.
2. Factors emphasized and teaching techniques.
3. Handwriting systems and materials.
4. Writing instruments and surfaces.

Based on the survey the authors feel that research in the area of teaching aids, letter form and orientation to it, body position, speed of handwriting, instructional sequence, writing instruments, evaluation tools, writing surfaces and media are necessary.

Herrick, V., and Erlebacher, The Evaluation of Legibility in Handwriting. In V. Herrick (Ed.), New Horizons for Research in Handwriting, Madison: University of Wisconsin Press, 1963.

This is an article that covers historic research in handwriting evaluation scales from 1910 to 1961. Thorndike, Ayres, Freeman, West and the Wisconsin Scales are included.

Rondinella, Oreste R., An Evaluation of Subjectivity of Elementary School Teachers in Grading Handwriting. Elementary English, 1963, 40, 531-532.

A study by Rodinella was conducted to see if subjective grading is occurring in handwriting. The factors that are considered in grading and how these facts are weighed by teachers were looked into. The study used 210 teachers familiar with the Freeman and Ayres scales to evaluate papers. Rodinella found that subjectivity does exist and many teachers do not rely on criteria established for scales. Many tend to rely on letter form, readability and neatness not using general quality as criteria. The relationship between teacher and expert grades showed substantial agreement. The author cites that instruction is less effective when grading is subjective and recommends the use of scales and training in their use.

Otto, W., Askov, E., and Cooper, C., Legibility Rating for Handwriting Samples: A Pragmatic Approach. Perceptual and Motor Skills, 1967, 25, 638.

The purpose of this study was to see if teachers, having prior set criteria for grading, could grade reliability without the use of a scale. Samples were used from 240, 4th and 6th grade students. Three judges rated the samples for legibility using the California Achievement Test and Wisconsin Scale. The papers were then judged by teachers without the use of scales. The study shows that the judges did about the same with or without a scale. It was concluded that once a background of criteria is established for evaluating handwriting, teachers do not need scales to make reliable judgements.

Bell, Mary Elizabeth, Evaluating the Quality of Handwriting. Education, 1969, 90, 126-129.

Bell has used the research of others to set up a way for teachers to judge the handwriting of pupils. Bell notes that importance needs to be given to selection of materials to evaluate, factors of legibility, goals, child development, noting improvement, and readability.

Andersen, Dan W., What Makes Writing Legible? The Elementary School Journal, 1969, 69, 365-368.

A study by Andersen investigated factors that make writing legible. Size, slant, uniformity of size and slant and legibility factors were studied. The findings on the relationships were: the more legible writing is, the larger it is; the more legible writing, the more uniform the slant; the larger the writing, the less uniform the size; and the more pronounced the slant, the more uniform the slant.

Groff, Patrick, The Future of Legibility. Elementary English, 1975, 52, 205-212.

Groff reviews the works of many authors citing the de-emphasis of handwriting legibility. It is this authors opinion the the lowering of legibility standards calls for an increase in technical advancement or the use of a functional shorthand to meet the needs of our students and society.

Helwig, J., Johns, J., Norman, J., and Cooper, J., The Measurement of Manuscript Letter Strokes. Journal of Applied Behavior Analysis, 1976, 9, 231-236.

Helwig presents a way to evaluate letter formation with the use of transparent overlays. Overlays consist of ellipse like boundaries that cover letters to be evaluated. The six criteria that can be judged in its confines are: strokes completeness, closed curves or circles, strokes intersect each successive stroke, letters are complete and horizontal crossing of t and f's intersect within the confines of the ellipse. Trained observers, as well as naive observers, were used to judge samples of students work. The study found that transparent overlays may be a significant tool for measuring individual letter formation. The time involved in using this tool increase with practice and the cost of the overlays are within reach of all school systems.

Sims, E., and Weisberg, P., Effects of Page Prompts on Beginning Handwriting Legibility. Journal of Educational Research, 1984, 77 (6) , 360-365.

This article reports on a study which compares legibility of letters using a start-ball prompt, both with and without pointer, to the method of using dotted letter tracing prompt and writing in a blank space. The study also addresses two evaluation procedures for assessing legibility of letters. The research showed that with the use of prompts the letter legibility improved when judged by an overlay. Differences were not as easily determined when teacher evaluation were used.

Ziviani, J., and Elkins, J., An Evaluation of Handwriting Performance. Educational Review, 1984, 36 (3) , 249-261.

The study used samples of 575 children grades 3-7 from four primary schools. The formations, spacing, spacing/range, alignment and size were evaluated. Four factors of legibility were judged by using transparent overlays as well as mathematical measurement for spacing. Speed of writing was judged by timing repetitions of words written. Results of the testing showed acceptable levels of reliability between judges. The lowest area was that of letter formation where individual bias may play a part. Speed of performance was the nost reliable measured in retesting. Legibility items fluctuated with alignment showing the lowest consistency. Younger children had more variability than the higher grade children on retesting. Validity showed moderate relationships among variables of legibility and criterion validity was done by a comparison of speed results of the test to Ayres and Groff. Legibility factors were compared to teacher ratings. Comparison among sex is included.

Groff, Patrick, Successful Remediation of Cursive Handwriting. Journal of Reading, Writing and Learning Disabilities, International, 1984, 1 (1) , 11-15.

Groff outlines a remediation plan for the transitional stage from manuscript to cursive for the handicapped writer. The first four stages to teach, is the idea of space, slant, alignment and size with the use of transparent overlay and lined paper. He feels that successful letter shape must follow these other factors. This method of remediation requires the teacher to closely monitor the pupils progress.

Graham, Steve, Reliability, Validity, and Utility of Three Handwriting Measurement Procedures. Journal of Education Research, 1986, 79 (6), 373-380.

A study was done to measure the reliability, validity and utility of three types of handwriting evaluation methods. The three methods analyzed are: correct/incorrect, holistic rating without models and a holistic rating with models. Results show of the three methods only the correct/incorrect method which uses the transparent overlay proved reliable in interater and intrarater measurement. This type of measurement proved not to be consistent to other forms of handwriting evaluation. Holistic scoring methods proved to be more consistent when compared to other measures. The reliability of the judges scores were low. The utility of all three methods were examined by the judges. Factors such as time, ease of grading, subjectivity and testing data showed the practical use of any method study was low.

Wolfson, B., The Educational Scene: The Teaching of Handwriting. Elementary English, 1962, 39, 55-59.

This article gives the results of two studies. The first was done to find out what the current practices are in teaching handwriting. This was a national survey of educators. The survey found 95% of teachers teach handwriting with both manuscript and cursive. The six factors of handwriting which are emphasized are: letter form, size, uniformity, spacing, alignment and speed, which was found to be a factor of less importance. The practice of handwriting instruction comes mainly from the commercial system which is used with only 1/3 of schools using a scale to evaluate handwriting.

The second study is on "Children's Perception of the Handwriting Task." This study details how children of different intelligence levels perceive their handwriting to that of adult judges. The findings were that bright and average childrens judgements were more in alignment than children in the dull group. All groups differed in choosing work that resembled their own, with the bright children having the most stable perception. This study questions the childs ability to judge handwriting using models.

Robin, A., Armel, S., and O'Leary, K., The Effects of Self-Instruction on Writing Deficiencies. Behavior Therapy, 1975, 6, 178-187.

This article details a study done to see if self-instruction using training in self produced verbalizations could improve the handwriting abilities. This was compared to a directed class control group that received no praise. Results showed that "self-instruction plus directed training proved more effective than directed training alone." Self-instruction with verbalization statistically proved significant but in practicality, was hard to teach and was looked on as bothersome by students.

Jones, J., Trap, J., and Cooper, J., Technical Report: Students' Self Recording of Manuscript Letter Strokes. Journal of Applied Behavior Analysis, 1977, 10, 509-514.

Transparent overlays were used to self train three groups of first grade students to correctly measure their manuscript letter strokes. Overlays were used with training sheets and approximately 80-120 minutes of instruction and practice. Students could judge their work with roughly 82% accuracy when compared to experimenters scores. This was visioned as a practical method when students can self check and get immediate feedback on letter strokes.

Lindsey, J., and Beck, F., Handwriting and the Classroom Experience: A Recapitulation. Pointer, 1984, 29 (1), 29-31.

This is a review of literature on letter formation designed to assist the classroom teacher. Classroom ideas are given to remediate problems in letter formation.

Rubin, N., and Henderson,S., Two Sides of the Same Coin: Variations in Teaching Methods and Failure to Learn to Write. Special Education: Forward Trends, 1982, 9 (4), 17-24.

Students inability to write may stem from methods of instruction, evaluation or lack of it in the classroom. A study was done to collect data from 82 teachers to find responses to methods used in handwriting instruction. Responses showed 73 teach handwriting as a separate subject. Presentations of handwriting vary from whole class to incidental instruction. Frequency and time spent in handwriting lessons varied significantly among the group surveyed. Rubin's study used 40 students, 20 visioned as poor writers and 20 as a control group. All students were selected by classroom teachers. Students were given 30-60 minutes of handwriting instructions a week as a separate subject. Testing was compared with teachers judgements to see if they were valid. These results showed considerable variability. Qualitative characters that affected writers ability were: legibility, formation and size, slant, spacing, speed and alignment. Conclusions call for systematic procedures to provide adequate remedial activities for students.

Graham, Steve, A Review of Handwriting Scales and Factors That Contribute to Variability in Handwriting Scores. Journal of School Psychology, 1986, 24 (1), 63-71.

This paper examines factors that effect the reliability and validity of handwriting evaluation. It also studied how useful the most common handwriting scales are to instruction. The results of the study show that although handwriting scales have limitations they can serve such purposes as a screening device. This can improve educators informal evaluations for students in need of instructional help. Recommendations were given to reduce measurement error in handwriting assessment.

Collins, F., Baer, G., Walls, N., and Jackson, M.S., The Development of a Behavioral Assessment Technique for Evaluating Gradual Change in Handwriting Performance. Behavioral Assessment, 1980, 2, 369-387.

This article represents three studies which evaluated methods of assessing gradual improvement in handwriting. The three methods studied were: forced-choice ranking, correct/incorrect scoring and a five point global scale of illegibilities in handwriting. The study showed that gradual change can be reliably detected by force/rank or a global procedure more readily than a correct/incorrect approach. The use of a force rank approach is more time consuming than the global scale, but it gives more specific feedback to stucents whose work is being evaluated. Force-rank criteria are: overall appearance, smoothness of stroke, closure, stroke level where a global approach is more subjective.

Peck, M., Askov, E., and Fairchild, S., Another Decade of Research in Handwriting: Progress and Prospect in the 1970's. Journal of Educational Research, 1980, 73 (5) , 283-298.

The purpose of this article is to review literature on handwriting for the decade of 1970-1980. Reference is given to current research, as well as to research of the past.

Graham, Steve, Measurement of Handwriting Skills: A Critical Review. Diagnostique, 1982, 8 , 32-40.

The author details a very thorough review of handwriting evaluation techniques. Scales, methods and informal procedures are reviewed along with findings of reliability, validity and utility in the use of them. He offers guidelines to improve the reliability and validity of handwriting tools. The evaluator should be trained in the use of the tool, identifying information should be removed, use more than one evaluator, be consistent in giving instruction to students and never rate handwriting when fatigued.

Summary

Much of the research of the early 1900's was spent on the development and comparison of handwriting scales. Dan Andersen's(1965) review points out that research in handwriting was popular in the first three decades of the 1900's. The other periods claiming WWII were relatively quiet up to 1950.

Finding a means of evaluating handwriting started with Professor Edward L. Thorndike in 1910. His handwriting scale consisted of handwriting samples ranked in quality from the level of 4-18 (with 4 being low). The samples were based on general merit which was concensus of competent judges. His research served as a basis for further study.

Ayres developed a handwriting scale shortly after Thorndike. The handwriting scale was made from actual student samples ranked from 10-90. The basis for writing evaluation scale was the amount of time necessary to read each sample.

The question of which scale is better was raised by Pitner(1914). He compared Thorndike's Scale to Ayres' Scale to find which scale yielded the most uniform results. In noting the ability of the scales to rank papers, they showed a significant similarity, but when compared by deviation, he showed the Thorndike scale gave more reliable and uniform results. Kelley (1915) questioned the accuracy of Pitner's findings. He felt that the way in which Pitner arrived at his measurements for both scales were not comparable for finding reliability or uniformity. When he did his own calculations, he found Ayres' scale to be slightly superior in uniformity.

Research that followed the development of the scales was done to test the usefulness and need for them. King(1912) used the Ayres scale to test

students samples with reference to practicability of the scale and results of age, sex and school variations. He showed that a single judge can use the scale with a fair amount of consistency but when the scale was used by eight judges it "was impossible to eliminate subjective factors." Manuel's(1915) study found that variability of measure tended to be greater without the use of a scale, but the scale method had no superiority over the other methods when used to rank papers. He also found that deviations among individuals or groups lessened as experience with the scale increased. Gray(1915) found that training is beneficial in reducing variability among scores.

Once again, the handwriting scales were compared. Starch(1919) felt it was necessary to compare Ayres' Scale to Thorndike's to see if perhaps a new means of evaluation was necessary. He evaluated the practicality of both the Ayres and Thorndike scales. Thorndike's scale is very awkward to handle and samples of different slants of handwriting are not available. The practicality of Ayres scale received little criticism. In rechecking values given to Thorndikes samples of handwriting significant differences were found among the ranking of handwriting. In the Ayres scale, marked inequalities were found in the distances between the steps.

Freeman(1915) began his work in handwriting evaluation by detailing steps which are necessary for more reliable use of scales. He defines slant, height, letter formation, quality of line, and appearance. He gives objective ways to measure them. His work is found in his Progress Record(1929) in which he defines the prior five areas of legibility and gives teachers specific ways in which to judge them.

Freeman(1959) developed a handwriting scale for the commercial use of Zaner Bloser. The scale was made using childrens handwriting samples gathered from around the nation. On the scale each grade level had

five samples, which were ranked from high to low. This scale was made to be representative of the quality of childrens handwriting throughout the country. Feldt(1962) conducted a study to find reliability data on the Freeman scale. He found that the relationship among judges scores on handwriting are low when compared to achievement tests. He does not feel that scales are useless. He suggested that teachers average scores from testing sessions and they should be provided with more training in handwriting evaluation.

Scales continued to be developed, Bezzi(1962) created a scale by adding to Freemans requirements of evaluating handwriting. She adds rate, overlap in quality among grades and cross-validation with a cursive scale, for a standardized measure of handwriting.

The question of how handwriting has progressed from the past to 1960 and the instructional process of teaching it was presented at a conference at the University of Wisconsin by Erlebacher, Herrick and Odaka. The study by Erlebacher and Herrick(1961) was done to make a valid assessment of the quality of writing today as it compares to 1912. The study revealed that handwriting does not show signs of deterioration. At the same conference, Herrick and Okada(1963) gave results of their survey to reveal the instructional process in the United States. Based on the survey they recommended research in many areas of handwriting including evaluation tools. Wolfson(1962) also studied current practices in handwriting instruction. She noted that "1/3 of schools use a scale to evaluate handwriting" and that the practice of handwriting instruction comes mainly from commercial systems.

A shift in research of handwriting evaluation occurred as research started to assess what the classroom teacher is actually basing his/her

evaluation on. Rodinella(1963) did a study to find out if subjective grading is occurring and what factors are considered in their grading. It was found that subjective grading is happening and many teachers do not rely on criteria set by scales. Teachers mainly look for letter form, readability and neatness. The relationship between teacher and expert grades are significant. This author offers the idea that instruction is less effective when based on subjective grading and recommended training in the use of scales. In Rodinella's study, teachers did have knowledge of the Freeman and Ayres Scale. Otto's(1967) study compared the legibility grades given by teachers with and without the Wisconsin scale. All teachers had a background of the criteria and the study proved that once a criteria is understood the use of the scale is not necessary for making reliable judgements.

Bell's(1969) review notes the importance of knowing factors to judge legibility. Ayres(1912) uses readability as a synonym for legibility. Freeman(1959) set up steps for evaluating legibility. Andersen(1969) looked deeper into relationships that makes writing legible. He found relationships between size, slant, uniformity and legibility. Groff reviews the works of many authors on the de-emphasis of handwriting legibility. It is his opinion that due to low standards of writing technical advances and a functional shorthand need to be developed.

The research of the late 1970's to the present covers such topics as: transparent overlays, students evaluating their own work, remedial uses of handwriting evaluation and reviews of literature in handwriting evaluation.

The use of transparent overlays to evaluate handwriting has received quite a bit of research. Helwig(1976) presents a way in which to evaluate

letter formation. The study used trained and naive observers. Transparent overlays were found to be a significant tool for measuring individual letter formation. Sims and Weisberg(1984) compared writing techniques and also assessed two evaluation techniques. They found that letter legibility improved when judged by overlays. Differences among letter forms were not as easily noted by teacher evaluation with no overlay. Ziviani's(1984) study used transparent overlays when trying to achieve acceptable levels of reliability evaluating factors of legibility. The factors considered were formation, spacing, spacing range, alignment and size. Groff(1984) used transparent overlays to check space, slant, alignment and size of letters. He says this evaluation method requires closely monitoring the pupils progress. Graham(1986) did a study to see how reliable, valid and usable three methods of evaluation are. He looked at the correct/incorrect method, holistic and holistic with models. The correct/incorrect method with an overlay proved to be the most reliable. The measures did not prove to be consistent with other handwriting methods. Holistic methods were more consistent but less reliable. The utility of all three methods were questionable.

The use of transparent overlays and student self checking of handwriting has also received various notice. In Wolfson's(1962) article she questioned a child's ability to judge his/her handwriting using models. Her work showed that bright children had the most stable perception of their own work. Robin(1975) tried to teach students to self instruct using self produced verbalization. The training in self verbalization proved statistically significant but in terms of practicality was hard to teach and bothersome to students. Jones(1977) and others trained students to judge letter strokes using a transparent overlay. With training, students

could accurately judge their work 82% of the time when compared to expert evaluations.

Handwriting assessment can be a way of detecting students in need of handwriting remediation. Several authors have reviewed literature or initiated studies on topics in this area. Lindsey's(1984) review of literature on letter formation gives ideas for remediation in the classroom. In her review she states, "Classroom teachers will find few commercial tests available for assessing handwriting."(p.30) Rubin(1982) felt that a students inability to write stems from the method of instruction, evaluation or lack of it in the classroom. Rubins study concluded that teachers need a systematic procedure to provide adequate remediation for students perceived as poor writers. Collins(1980) evaluated methods of handwriting assessment to see how effective they were at noting gradual change. The study found that gradual changes were detected by forced rank or a global approach easier than by correct/incorrect criteria. The author also noted that forced rank evaluation is more time consuming but lends itself well to remediation by giving specific feedback. Graham(1986) found that although scales have limitations they can be used as screening devices for students in need of instructional help. They can also improve educators informal evaluation of handwriting.

Authors like Peck, Askov, Fairchild and Graham continue to review research articles on handwriting evaluation. Pecks(1980) review hopes to stimulate more research. Grahams(1982) critical review gives guide lines to improve reliability and validity of research in handwriting assessment. He encourages readers to remember that "handwriting is a subskill of written composition it is not an end unto itself, but rather a means to an end. Assessment should be geared toward insuring that handwriting is produced

with maximum efficiency."

Conclusion

Many varied views on the evaluation of handwriting exist. There seems to be no best single method. It is important to be aware of evaluation techniques and tools that can be used to judge students handwriting skills.

The earliest research in handwriting was in the making of handwriting scales. Each tool added some new dimension of objectivity to evaluation, yet still received various criticisms. Training, practice and familiarizing educators in the use of scales can reduce the variability in their use.

Subjectivity among evaluators of handwriting does occur. There does seem to be a significant relationship between rankings and grades given by teachers when compared to experts. This may be due to the fact that once an evaluation criteria is established the scales use is no longer necessary.

Research keeps adding and redefining legibility and criteria to judge it. Teachers may look at size, slant and uniformity when judging samples. Without criteria for legibility handwriting instructors may have to turn to technological advances and functional shorthand to replace this training.

Transparent overlays seem to be an effective way to evaluate letter form. It has proven to be quite reliable when judging legibility factors. This tool will call for teachers to closely monitor pupils progress.

With transparent overlays and training students can self monitor their handwriting progress. Some research did find that a students perception of their own work is inaccurate. The use of transparent overlays to help students judge letter strokes is fairly accurate.

Handwriting instruction is less effective when based on subjective grading. Teachers need ways to remediate by noticing gradual changes and by using methods that give specific feedback.

Existing research points to the need for constant review and study in the area of handwriting evaluation.

Recommendations

1. Choose a handwriting evaluation scale, method or tool and use it consistently.

2. Students should self check their handwriting by using transparent overlays to evaluate daily progress.

3. Use your evaluation plan as a tool which lends itself to remediation.

4. Research in handwriting evaluation should continue.

Bibliography

Andersen, Dan W., Handwriting Research: Movement and Quality. Elementary English, 1965, 42, 45-53. (6, 14)

Andersen, Dan W., What Makes Writing Legible? The Elementary School Journal, 1969, 69, 365-368. (9, 17)

Ayres, Leonard P., A Scale for Measuring the Quality of Handwriting of School Children. No. 113, Russell Sage Foundation. 1912. (2, 3, 4, 14, 17)

Bell, Mary Elizabeth, Evaluating the Quality of Handwriting. Education, 1969, 90, 126-129. (9, 17)

Bezzi, Rocky, Standardized Manuscript Scale for Grades 1, 2, and 3. The Journal of Educational Research, 1962, 55, No. 7, 339-340. (7, 16)

Collins, J., Baer, G., Walls, N., and Jackson, M.S., The Development of a Behavioral Assessment Technique for Evaluation Gradual Change in Handwriting Performance. Behavioral Assessment, 1980, 2, 369-387. (13, 19)

Erlebacher, Adrienne, and Herrick, Virgil E., Quality of Handwriting Today and Yesterday. Elementary School Journal, 1961, 62, 89-93. (7, 16)

Feldt, Leonard S., The Reliability of Measures of Handwriting Quality. Journal of Educational Psychology, 1962, 53, No. 6, 288-292. (7, 16)

Freeman, Frank, An Analytical Scale for Judging Handwriting. The Elementary School Journal, 1915, 15, 432-441. (2, 3, 6, 15)

Freeman, Frank Nugent, The Zaner-Bloser Progress Record and Scale in Handwriting. Zaner-Bloser Co., Columbus, Ohio, 1929. (6, 15)

Freeman, Frank N., A New Handwriting Scale. Elementary School Journal, 1959, 59, 218-221. (7, 15, 17)

Graham, Steve, Measurement of Handwriting Skills: A Critical Review. Diagnostique, 1982, 8, 32-42. (13, 19)

Graham, Steve, A Review of Handwriting Scales and Factors that Contribute to Variability in Handwriting Scores. Journal of School Psychology, 1986, 24 (1), 63-71. (2, 12, 18)

Graham, Steve, Reliability Validity, and Utility of Three Handwriting Measurement Procedures. Journal of Educational Research, 1986, 79 (6), 373-380. (11, 19)

Gray, Truman C., The Training of Judgement in the Use of the Ayres Scale for Handwriting. Journal of Educational Psychology, 1915, 6, 85-98. (5, 15)

Groff, Patrick, The Future of Legibility. Elementary English, 1975, 52, 205-212. (9, 17)

Groff, Patrick, Successful Remediation of Cursive Handwriting. Journal of Reading, Writing, and Learning Disabilities, International, 1984, 1 (1), 11-15. (10, 18)

Helwig, J., Johns, J., Norman, J., and Cooper, J., The Measurement of Manuscript Letter Strokes. Journal of Applied Behavior Analysis, 1976, 9, 231-236. (9, 17, 19)

Herrick, V., and Erlebacher, The Evaluation of Legibility in Handwriting. In V. Herrick (Ed.), New Horizons for Research in Handwriting, Madison: University of Wisconsin Press, 1963. (8, 16)

Jones, J., Trap, J., and Cooper, J., Technical Report: Students' Self-Recording of Manuscript Letter Strokes. Journal of Applied Behavior Analysis, 1977, 10, 509-514. (12, 18)

Kelley, Truman Lee, Comparable Measures. Journal of Educational Psychology, 1915, 6, 589-595. (4, 14)

King, I., and Johnson, H., The Writing Abilities of the Elementary and Grammar School Pupil of a City School System Measured by the Ayres Scale. Journal of Educational Psychology, 1912, 3, 514-520. (5, 14, 15)

Lindsey, J., and Beck, F., Handwriting and the Classroom Experience: A Recapitulation. Pointer, 1984, 29 (1), 29-31. (12, 19)

Manuel, Herschel T., The Use of an Objective Scale for Grading Handwriting. The Elementary School Journal, 1915, 15, 269-278. (5, 15)

Otto, W., Askov, E., and Cooper, C., Legibility Rating for Handwriting Samples: A Pragmatic Approach. Perceptual and Motor Skills, 1967, 25, 638. (8, 16)

Peck, M., Askov, E., and Fairchild, S., Another Decade of Research in Handwriting: Progress and Prospect in the 1970's. Journal of Educational Research, 1980, 13 (5), 283-298. (13, 19)

Pitner, Rudolf, A Comparison of the Ayres and Thorndike Handwriting Scales. Journal of Educational Psychology, 1914, 5, 525-536. (4, 14)

Robin, A., Armel, S., and O'Leary, K., The Effects of Self-Instruction on Writing Deficiencies. Behavior Therapy, 1975, 6, 178-187. (11, 18)

Rondinella, Oreste R., An Evaluation of Subjectivity of Elementary School Teachers in Grading Handwriting. Elementary English, 1963, 40, 531-532. (8, 17)

Rubin, N., and Henderson, S., Two Sides of the Same Coin: Variations in Teaching Methods and Failure to Learn to Write. Special Education: Forward Trends, 1982, 9 (4), 17-24. (12, 19)

Sims, E., and Weisberg, P., Effects of Page prompts on Beginning Handwriting Legibility. Journal of Educational Research, 1984, 77 (6), 360-365. (10, 18)

Starch, Daniel, Educational Research and Statistics: A Scale for Measuring Handwriting. School and Society, 1919, 9, 155-158. (6, 15)

Thorndike, Edward L., Handwriting, Teachers College, Columbia University, New York, 1912. (1, 4, 14)

Wolfson, B., The Educational Scene: The Teaching of Handwriting. Elementary English, 1962, 39, 55-59. (11, 16, 18)

Ziviani, J., and Elkins, J., An Evaluation of Handwriting Performance. Educational Review, 1984, 36 (3), 249-261. (10, 18)

Student Background

Name: Rene Formsma

Address: 53910 Juanita Dr.
Elkhart, IN 46514

Student Number: 375-66-4756

Phone: (219) 262-3380

Age: 28

Marital Status: Married

Children: Paige, 10 months
Derek, 3 years

Education:

Currently, Graduate Student at I.U.S.B.

B.S., Central Michigan University in 1981.

Employment:

Currently, Second Grade Teacher with Elkhart Community Schools at Mary Feeser.

Other Experiences, First Grade Teacher with Elkhart, Kansas Schools. Seventh and Eighth Grade Art, Social Studies and American History Teacher with Syracuse, Kansas Schools.

Reason:

Elkhart Community Schools teach handwriting at the second grade level as an independent subject. I want to be able to use an objectice approach when evaluating this skill for my students. This can improve my instruction of handwriting skills and be beneficial to students.